On a Conjecture of Gohberg and Rodman

Jim Hartman*
The College of Wooster
Wooster, Ohio 44691

Submitted by Leiba Rodman

ABSTRACT

Gohberg and Rodman in *Linear Algebra Appl.* 76:85–120 (1986) give an algorithm to construct a sequence of matrices supposedly satisfying certain conditions related to Conjecture 7.2 in that paper. Unfortunately, this construction fails to satisfy those conditions. We give a different sequence having the desired properties along with verifying the conjecture for another special case. Finally, we show that the conjecture, in its generality, is false.

We adopt the notation found in [2]. For a subspace N in $\mathbb{C}^n$ ($n \geqslant 2$) we will let P_N denote the orthogonal projection onto N. For two subspaces M and N in $\mathbb{C}^n$ we let $\theta(M,N) = \|P_M - P_N\|$, which is a metric on the set of all subspaces of $\mathbb{C}^n$ with $\theta(M,N) \leqslant 1$ for all M and N. We note that if $\dim M \neq \dim N$ then $\theta(M,N) = 1$.

If $A : \mathbb{C}^n \to \mathbb{C}^n$ is a linear transformation, we let $\operatorname{Inv} A$ denote the lattice of subspaces invariant under A (i.e., a subspace M is in $\operatorname{Inv} A$ if and only if for all $m \in M$ we have $Am \in M$). We let $\mathscr{B}(\mathbb{C}^n)$ denote the space of all linear transformations on $\mathbb{C}^n$. If M is a subspace of $\mathbb{C}^n$ and $A \in \mathscr{B}(\mathbb{C}^n)$, we let $d(M, \operatorname{Inv} A) = \inf\{\theta(M,N) : N \in \operatorname{Inv} A\}$. For $A, B \in \mathscr{B}(\mathbb{C}^n)$ we define the distance between $\operatorname{Inv} A$ and $\operatorname{Inv} B$ by

$$\operatorname{dist}(\operatorname{Inv} A, \operatorname{Inv} B)$$
$$= \max\{\sup\{d(M, \operatorname{Inv} B) : M \in \operatorname{Inv} A\}, \sup\{d(N, \operatorname{Inv} A) : N \in \operatorname{Inv} B\}\}.$$

*The author would like to thank the Mathematics Department at Dalhousie University, where this work was completed while he was visiting there. The author would also like to thank Ben Mathes for many helpful discussions.

655 Avenue of the Americas, New York, NY 10010 0024-3795/90/$3.50

For $A \in \mathscr{B}(\mathbb{C}^n)$ we let $\sigma(A) = \{\lambda \in \mathbb{C} : A - \lambda I$ is not invertible$\}$, that is, the set of eigenvalues of A. If $\sigma(A) = \{\lambda_1, \dots, \lambda_s\}$ then the geometric multiplicity of λ_i is denoted by $r_i = \dim \operatorname{Ker}(A - \lambda_i I)$. We may assume that $r_s \leqslant \cdots \leqslant r_1$. Then in the Jordan canonical form of A there are r_i blocks corresponding to the eigenvalue λ_i. We let $m_{i1} \geqslant m_{i2} \geqslant \cdots \geqslant m_{ir_i}$ be the sizes of those blocks, and then we define

$$\Omega_A = \{s; r_1, r_2, \dots, r_s; m_{11}, m_{12}, \dots, m_{1r_1}; m_{21}, \dots, m_{2r_2}; \dots; m_{s1}, \dots, m_{sr_s}\}.$$

Here we can assume that if $r_i = r_{i+1}$ then

$$\sum_{j=1}^{r_i} m_{ij} \geqslant \sum_{j=1}^{r_i} m_{(i+1)j},$$

and if equality holds here, then

$$\sum_{j=1}^{k} m_{ij} \geqslant \sum_{j=1}^{k} m_{(i+1)j} \qquad \text{for} \quad k = 1, 2, \dots, r_{i-1}.$$

Hence Ω_A describes the block Jordan structure without regard to the actual values of the eigenvalues. For Δ a subset of $\{1, 2, \dots, s\}$ we let $k_j(\Omega_A, \Delta) = \sum_{i \in \Delta} m_{ij}$ for $j = 1, 2, \dots,$ where $m_{ij} = 0$ for $j > r_i$.

We let $P(\Omega_A)$ denote the set of Ω' which satisfy the following condition, where $\Omega' = \{s'; r'_1, \dots, r'_{s'}; m'_{11}, \dots, m'_{1r'_1}; \dots; m'_{s'1}, \dots, m'_{s'r'_{s'}}\}$:

There is a partition of $\{1, 2, \dots, s'\}$ into nonempty, disjoint sets $\Delta_1, \dots, \Delta_s$ such that

$$\sum_{j=1}^{t} k_j(\Omega_A, \{i\}) \leqslant \sum_{j=1}^{t} k_j(\Omega', \Delta_i) \qquad \text{for} \quad t = 1, 2, 3, \dots \text{ and } i = 1, 2, \dots, s$$

and such that equality holds when $t = \infty$.

By Theorem 15.10.2 of [2], there exists a sequence $\{A_m\}_{m=1}^{\infty}$ in $\mathscr{B}(\mathbb{C}^n)$ with $\|A_m - A\| \to 0$ as $m \to \infty$ and $\Omega_{A_m} = \Omega'$ if and only if Ω' satisfies the condition above.

We let

$$J_n = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & & \vdots \\ 0 & \cdot & \cdot & \cdot & 0 & 1 \\ 0 & \cdot & \cdot & \cdot & 0 & 0 \end{bmatrix}$$

be an $n \times n$ matrix, and note that $\Omega_{J_n} = \{1; 1; n\}$. Hence by the work above, $\Omega \in P(\Omega_A)$ if and only if $\Omega = \{s; 1, \ldots, 1; m_1; \ldots; m_s\}$. In Gohberg and Rodman [1] an algorithm is given which is supposed to yield a sequence $\{B_m\}_{m=1}^{\infty}$ such that $\operatorname{dist}(\operatorname{Inv} B_m, \operatorname{Inv} J_n)/\|B_m - J_n\| \to \infty$ as $m \to \infty$ and $\Omega_{B_m} = \Omega$ with $\Omega \neq \Omega_{J_n}$. However, the construction given there does not work. Consider the case when $n = 3$. For $\Omega = \{2; 1, 1; 2; 1\}$ the algorithm gives

$$B_\varepsilon = \begin{bmatrix} 0 & 1 & 0 \\ \varepsilon & 0 & 1 \\ 0 & 0 & \sqrt{\varepsilon} \end{bmatrix}.$$

It is easy to see that $\|B_\varepsilon - J_3\| = \sqrt{\varepsilon}$ when $\varepsilon \leqslant 1$. If $\mathbf{e}_i$ denotes the ith coordinate vector in $\mathbb{C}^n$, then

$$\operatorname{Inv} J_3 = \{\{\mathbf{0}\}, \operatorname{sp}\{\mathbf{e}_1\}, \operatorname{sp}\{\mathbf{e}_1, \mathbf{e}_2\}, \mathbb{C}^3\}$$

and

$$\operatorname{Inv} B_\varepsilon = \{\{0\}, \operatorname{sp}\{\mathbf{f}_1\}, \operatorname{sp}\{\mathbf{f}_2\}, \operatorname{sp}\{\mathbf{f}_1, \mathbf{f}_2\}, \operatorname{sp}\{\mathbf{f}_1, \mathbf{f}_3\}, \mathbb{C}^3\},$$

where

$$\mathbf{f}_1 = \mathbf{e}_1 + \sqrt{\varepsilon}\, \mathbf{e}_2, \qquad \mathbf{f}_2 = \mathbf{e}_1 - \sqrt{\varepsilon}\, \mathbf{e}_2, \quad \text{and} \quad \mathbf{f}_3 = \mathbf{e}_2 + 2\sqrt{\varepsilon}\, \mathbf{e}_3.$$

One can check easily that $\operatorname{dist}(\operatorname{Inv} B_\varepsilon, \operatorname{Inv} J_3) \leqslant c\sqrt{\varepsilon}$ for some constant c not depending on ε by just finding $\theta(M, N)$ for all M in $\operatorname{Inv} J_3$ and all N in $\operatorname{Inv} B_\varepsilon$. Hence, $\operatorname{dist}(\operatorname{Inv} B_\varepsilon, \operatorname{Inv} J_3)/\|B_\varepsilon - J_3\| \leqslant c$ for all ε. For the sequence one should use $\varepsilon = 1/m$.

However the authors are correct in their assertion that such a sequence can be found, as Theorem 2 indicates. We first prove the following lemma.

LEMMA 1. *For $k > 1$ let $m_1, \ldots, m_k$ be positive integers. Then there are k distinct complex numbers $c_1, \ldots, c_k$ such that $c_k \neq 0$ and*

$$\sum_{i=1}^{k} c_i m_i = 0.$$

Proof. We use induction on k. Certainly this is true for $k = 2$ by taking $c_2 = 1$, and $c_1 = -c_2 m_2 / m_1$. Now assume $k = n+1$. By the induction hypothesis there exist distinct complex numbers $d_1, d_2, \ldots, d_n$ such that $\sum_{i=1}^{n} d_i m_i = 0$ and $d_n \neq 0$. If $S = \sum_{i=1}^{k} m_i$, pick a complex number $t \neq 0$ such that $t \neq d_i m_k / S$ for $i = 1, 2, \ldots, k-1 = n$. Then we will have

$$\sum_{i=1}^{n} (d_i - t) m_i + \left(t \sum_{j=1}^{n} \frac{m_j}{m_k} \right) m_k = 0.$$

Now for $i = 1, \ldots, k-1$ we let $c_i = d_i - t$ and $c_k = t \sum_{j=1}^{n} m_j / m_k$. Now certainly $c_1, \ldots, c_{k-1}$ are distinct. If $c_i = c_k$ for $i < k$, then we would have $t = d_i m_k / S$, which contradicts our choice of t. ■

We now have the following theorem.

THEOREM 2. *Let $J = J_n$, and let $\Omega = \{s; 1, \ldots, 1; m_1; \ldots; m_s\}$ with $s > 1$. Then there exists a sequence $\{B_m\}_{m=1}^{\infty}$ in $B(\mathbb{C}^n)$ such that $\Omega_{B_m} = \Omega$ and* $\operatorname{dist}(\operatorname{Inv} J, \operatorname{Inv} B_m) / \|J - B_m\| \to \infty$ *as $m \to \infty$.*

Proof. We first apply Lemma 1 to get distinct complex numbers $c_1, c_2, \ldots, c_s$ such that $c_s \neq 0$ and $\sum_{i=1}^{s} c_i m_i = 0$. Let $\lambda_i = c_i / (c_s \sqrt{m})$. Then $\lambda_1, \lambda_2, \ldots, \lambda_s$ are distinct, $\lambda_s = 1/\sqrt{m}$, and $\sum_{i=1}^{s} m_i \lambda_i = 0$. We now define

$$p(z) = \prod_{i=1}^{s} (z - \lambda_i)^{m_i}.$$

Then $p(z) = z^n - (\sum_{i=1}^{s} m_i \lambda_i) z^{n-1} - b_{m(n-2)} z^{n-2} - \cdots - b_{m1} z - b_{m0}$. We let

$$B_m = J + \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ b_{m0} & \cdots & b_{m(n-2)} & 0 \end{bmatrix}$$

be the companion matrix for $p(z)$. Then we will have $\Omega_{B_m} = \Omega$ by Theorem 3.3.14 of [3] and $\sigma(B_m) = \{\lambda_1, \dots, \lambda_s\}$. An eigenvector corresponding to $\lambda_s = 1/\sqrt{m} \equiv b_m$ is given by $\mathbf{f}_s = \sum_{i=1}^{n}(1/\sqrt{m})^{i-1}\mathbf{e}_i$. Now $N = \mathrm{sp}\{\mathbf{f}_s\}$ is in $\mathrm{Inv}\, B_m$, and $M = \mathrm{sp}\{\mathbf{e}_1\}$ is the only subspace of dimension 1 in $\mathrm{Inv}\, J$. Hence $d(N, \mathrm{Inv}\, J) = \theta(M, N)$. Now

$$P_M = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix},$$

and if $I_m = (m-1)/(m - m^{-(n-1)})$ then

$$P_N = I_m \begin{bmatrix} 1 & b_m & b_m^2 & \cdots \\ b_m & b_m^2 & b_m^3 & \cdots \\ b_m^2 & b_m^3 & b_m^4 & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}.$$

Hence $\theta(M, N) = \|P_N - P_M\| \geqslant \|(P_N - P_M)\mathbf{e}_1\| \geqslant I_m/\sqrt{m}$. This shows that

$$\mathrm{dist}(\mathrm{Inv}\, B_m, \mathrm{Inv}\, J) \geqslant I_m/\sqrt{m}.$$

We also have that $\|J - B_m\| \leqslant \sqrt{n}\, \max\{|b_{mi}| : i = 0, 1, \dots, n-2\}$ and note that b_{mi}, which is a coefficient of $p(z)$, is the linear combination of the product of two or more of the eigenvalues of B_m. Since the eigenvalues of B_m are $\lambda_i = c_i/(c_s\sqrt{m})$, we have that $\sqrt{n}\,|b_{mi}| \leqslant k/m$, where k depends only on n and $c_1, c_2, \dots, c_s$. As none of these depend on m, k does not depend on m. Thus

$$\frac{\mathrm{dist}(\mathrm{Inv}\, B_m, \mathrm{Inv}\, J)}{\|J - B_m\|} \geqslant \frac{I_m/\sqrt{m}}{k/m} = \frac{I_m\sqrt{m}}{k} \to \infty \quad \text{as} \quad m \to \infty. \qquad \blacksquare$$

Conjecture 7.2 of [1] (Conjecture 16.7.1 of [2]) suggests that for any $A \in \mathscr{B}(\mathbb{C}^n)$ and any $\Omega \in P(\Omega_A)$ with $\Omega \neq \Omega_A$ there is a sequence $\{B_m\}_{m=1}^{\infty}$ in $\mathscr{B}(\mathbb{C}^n)$ such that $\Omega_{B_m} = \Omega$ and $\mathrm{dist}(\mathrm{Inv} A, \mathrm{Inv}\, B_m)/\|A - B_m\| \to \infty$ as $m \to \infty$. Theorem 2 implies that this conjecture is true for any A that has

$\dim \operatorname{Ker}(A - \lambda I) = 1$ for all $\lambda \in \sigma(A)$. We now extend this result to any A for which $\operatorname{Ker}(A - \lambda I) = R_\lambda(A)$ for all $\lambda \in \sigma(A)$, where

$$R_\lambda(A) = \left\{ \mathbf{f} \in \mathbb{C}^n : (A - \lambda I)^k \mathbf{f} = \mathbf{0} \text{ for some positive integer } k \right\}.$$

We note that it is sufficient to consider only an A which has one eigenvalue, and with no loss of generality we may assume $A = 0$.

THEOREM 3. *Let $A = 0$, and let $\Omega \in P(\Omega_A)$ with $\Omega \neq \Omega_A$. Then there is a sequence $\{B_m\}_{m=1}^{\infty}$ in $\mathscr{B}(\mathbb{C}^n)$ such that $\Omega_{B_m} = \Omega$ and* $\operatorname{dist}(\operatorname{Inv} J, \operatorname{Inv} B_m) / \|A - B_m\| \to \infty$ *as $m \to \infty$.*

Proof. We note that $\Omega_A = \{1; n; 1, 1, \dots, 1\}$. Thus $\Omega \in P(\Omega_A)$ for any block Jordan structure

$$Q = \{s; r_1, r_2, \dots, r_s; m_{11}, \dots, m_{1r_1}; m_{21}, \dots, m_{2r_2}; \dots; m_{s1}, \dots, m_{sr_s}\}.$$

Let $\Omega \in P(\Omega_A)$ with $\Omega \neq \Omega_A$, and let T be a matrix in Jordan canonical form with $\Omega_T = \Omega$. If $B_m = (1/m)T$, then $\|B_m - A\| = \|T\|/m \to 0$ as $m \to \infty$ and $\Omega_{B_m} = \Omega$. Also $\operatorname{Inv}(B_m) = \operatorname{Inv}(T)$ for any m. Since $\Omega \neq \Omega_A$, we have that T has either two distinct eigenvalues or one eigenvalue with a block of size > 1.

If T has at least two eigenvalues, let $\mathbf{g}_1$ and $\mathbf{g}_2$ be two eigenvectors corresponding to two eigenvalues with $\|\mathbf{g}_i\| = 1$ for $i = 1, 2$. Note that $\langle \mathbf{g}_1, \mathbf{g}_2 \rangle = 0$, since T is in Jordan canonical form. Let $\mathbf{g} = \mathbf{g}_1 + \mathbf{g}_2$ and $M = \operatorname{sp}\{\mathbf{g}\}$. Then $M \in \operatorname{Inv} A$. If $N = \operatorname{sp}\{\mathbf{f}\}$ is any one-dimensional T invariant subspace, then $\mathbf{f} \perp \mathbf{g}_1$ or $\mathbf{f} \perp \mathbf{g}_2$. Assume with no loss of generality that $\mathbf{f} \perp \mathbf{g}_1$. Then we have

$$\begin{aligned} \|P_N - P_M\| &\geq \|(P_N - P_M)\mathbf{g}_1\| \\ &= \left\| 0 - \tfrac{1}{2}\langle \mathbf{g}_1, \mathbf{g}_1 \rangle \mathbf{g}_1 - \tfrac{1}{2}\langle \mathbf{g}_1, \mathbf{g}_2 \rangle \mathbf{g}_2 \right\| \\ &= \tfrac{1}{2}\|\mathbf{g}_1\| = \tfrac{1}{2}. \end{aligned}$$

Hence we will have $\theta(M, N) \geq \frac{1}{2}$ for any one-dimensional $N \in \operatorname{Inv} B_m$, giving $\operatorname{dist}(\operatorname{Inv} A, \operatorname{Inv} B_m) \geq d(M, \operatorname{Inv} B_m) \geq \frac{1}{2}$. Thus we have the desired result when there are two or more eigenvalues.

If T has only one eigenvalue, then it must have a Jordan block of size > 1. Let $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_k$ be a Jordan basis for that block of T with $\|\mathbf{f}_i\| = 1$ for

$i=1,\dots,k$. Now $M=\mathrm{sp}\{\mathbf{f}_1+\mathbf{f}_2\}\in \mathrm{Inv}\,A$. If $N=\mathrm{sp}\{\mathbf{f}\}\in \mathrm{Inv}\,T$ then $\mathbf{f}=\lambda \mathbf{f}_1+\mathbf{g}$, where $\mathbf{g}\perp \mathbf{f}_1$ and $\mathbf{g}\perp \mathbf{f}_2$. Thus we have

$$\begin{aligned}\|P_N-P_M\| &\geqslant \|(P_N-P_M)\mathbf{f}_2\|\\ &=\|\mathbf{0}-\tfrac{1}{2}\langle \mathbf{f}_2,\mathbf{f}_1\rangle \mathbf{f}_1-\tfrac{1}{2}\langle \mathbf{f}_2,\mathbf{f}_2\rangle \mathbf{f}_2\|\\ &=\tfrac{1}{2}\|\mathbf{f}_2\|=\tfrac{1}{2}.\end{aligned}$$

Again we have $\theta(M,N)\geqslant \frac{1}{2}$ for any one-dimensional $N\in \mathrm{Inv}\,B_m$, giving $\mathrm{dist}(\mathrm{Inv}\,A,\mathrm{Inv}\,B_m)\geqslant d(M,\mathrm{Inv}\,B_m)\geqslant \frac{1}{2}$.

Hence we have the desired result. ■

Thus to prove that the conjecture is true, the only remaining case to consider is when $A=J_{m_1}(0)\oplus J_{m_2}(0)\oplus\cdots\oplus J_{m_k}(0)$. Possibly the simplest case of this is for $A=J_2(0)\oplus J_1(0)$. As Theorem 6 indicates, the conjecture is not true in this case. Before proving Theorem 6 we establish the following lemmas.

LEMMA 4. *Let* $\mathbf{x}_1,\mathbf{x}_2\in \mathbb{C}^n$ *be linearly independent with* $1\leqslant \|\mathbf{x}_i\|\leqslant 2$ *for* $i=1,2$ *and* $|\langle \mathbf{x}_1,\mathbf{x}_2\rangle|\leqslant r<1$. *Let* $\mathbf{y}_1,\mathbf{y}_2$ *satisfy the same conditions along with having* $\|\mathbf{x}_i-\mathbf{y}_i\|\leqslant \varepsilon<1$ *for* $i=1,2$. *If* $M=\mathrm{sp}\{\mathbf{x}_1,\mathbf{x}_2\}$ *and* $N=\mathrm{sp}\{\mathbf{y}_1,\mathbf{y}_2\}$ *then* $\theta(M,N)\leqslant 272\varepsilon/(1-r^2)^2$.

Proof. Let $T=[\mathbf{x}_1\ \ \mathbf{x}_2]$ (i.e., the ith column of T is $\mathbf{x}_i$) and $S=[\mathbf{y}_1\ \ \mathbf{y}_2]$. Then $P_M=T(T^*T)^{-1}T^*$ and $P_N=S(S^*S)^{-1}S^*$, so we have

$$\begin{aligned}\theta(M,N)=\|&S(S^*S)^{-1}S^*-T(S^*S)^{-1}S^*+T(S^*S)^{-1}S^*\\ &-T(S^*S)^{-1}T^*+T(S^*S)^{-1}T^*-T(T^*T)^{-1}T^*\|\end{aligned}$$

By the triangle inequality

$$\begin{aligned}\theta(M,N)\leqslant{}&\|S-T\|\,\|(S^*S)^{-1}\|\,\|S^*\|+\|T\|\,\|(S^*S)^{-1}\|\,\|S^*-T^*\|\\ &+\|T\|\,\|(S^*S)^{-1}-(T^*T)^{-1}\|\,\|T^*\|.\end{aligned}$$

Now we have the following bounds on the norms above. Since $S-T=[\mathbf{x}_1-\mathbf{y}_1\ \ \mathbf{x}_2-\mathbf{y}_2]$, we have $\|S-T\|\leqslant\sqrt{2}\,\max\{\|\mathbf{x}_i-\mathbf{y}_i\|: i=1,2\}\leqslant\varepsilon\sqrt{2}$, $\|S^*\|$

$= \|S\| \leqslant \sqrt{2}\max\{\|\mathbf{y}_i\| : i = 1,2\} \leqslant 2\sqrt{2}$, and $\|T^*\| = \|T\| \leqslant \sqrt{2}\max\{\|\mathbf{x}_i\| : i = 1,2\} \leqslant 2\sqrt{2}$. Now,

$$T^*T = \begin{bmatrix} \|\mathbf{x}_1\|^2 & \langle \mathbf{x}_1, \mathbf{x}_2\rangle^* \\ \langle \mathbf{x}_1, \mathbf{x}_2\rangle & \|\mathbf{x}_2\|^2 \end{bmatrix},$$

so that for $t = (\|\mathbf{x}_1\|^2\|\mathbf{x}_2\|^2 - |\langle \mathbf{x}_1, \mathbf{x}_2\rangle|^2)^{-1}$ we have

$$(T^*T)^{-1} = t\begin{bmatrix} \|\mathbf{x}_2\|^2 & -\langle \mathbf{x}_1, \mathbf{x}_2\rangle^* \\ -\langle \mathbf{x}_1, \mathbf{x}_2\rangle & \|\mathbf{x}_1\|^2 \end{bmatrix},$$

and since $|\langle \mathbf{x}_1, \mathbf{x}_2\rangle| \leqslant r < 1 \leqslant \|\mathbf{x}_i\|$ for each $i = 1,2$, we have $\|(T^*T)^{-1}\| \leqslant t\max\{2\|\mathbf{x}_i\|^2 : i = 1,2\} \leqslant 2/(1-r^2)$. Similarly we have

$$\|(S^*S)^{-1}\| \leqslant 2/(1-r^2).$$

Using what we have above,

$$\begin{aligned} \left\|(S^*S)^{-1} - (T^*T)^{-1}\right\| &= \left\|(T^*T)^{-1}(T^*T - S^*S)(S^*S)^{-1}\right\| \\ &\leqslant \left\|(T^*T)^{-1}\right\| \|T^*T - T^*S + T^*S - S^*S\| \left\|(S^*S)^{-1}\right\| \\ &\leqslant \frac{4}{(1-r^2)^2}(\|T^*\|\,\|T-S\| + \|T^* - S^*\|\,\|S\|) \\ &\leqslant \frac{32\varepsilon}{(1-r^2)^2}. \end{aligned}$$

Using all these inequalities, we have

$$\begin{aligned} \theta(M,N) &\leqslant \varepsilon\sqrt{2}\,\frac{2}{1-r^2}2\sqrt{2} + 2\sqrt{2}\,\frac{2}{1-r^2}\varepsilon\sqrt{2} + 2\sqrt{2}\,\frac{32\varepsilon}{(1-r^2)^2}2\sqrt{2} \\ &\leqslant \frac{8\varepsilon}{(1-r^2)^2} + \frac{8\varepsilon}{(1-r^2)^2} + \frac{256\varepsilon}{(1-r^2)^2} = \frac{272\varepsilon}{(1-r^2)^2}. \end{aligned}$$ ∎

LEMMA 5. *Let* $1 \leqslant \|x_1\| \leqslant 2$ *and* $1 \leqslant \|y_1\| \leqslant 2$ *with* $\|x_1 - y_1\| \leqslant \varepsilon$. *If* $M = \mathrm{sp}\{x_1\}$ *and* $N = \mathrm{sp}\{y_1\}$ *then* $\theta(M, N) \leqslant 16\varepsilon$.

Proof. We let $T = [x_1]$ and $S = [y_1]$. Then $\|S^*\| = \|S\| \leqslant 2$ and $\|T^*\| = \|T\| \leqslant 2$. Also $\|S - T\| = \|x_1 - y_1\| \leqslant \varepsilon$, $(T^*T)^{-1} = 1/\|x_1\|^2$, and $(S^*S)^{-1} = 1/\|y_1\|^2$. Now using inequalities from the previous lemma, we get

$$\theta(M,N) \leqslant \varepsilon\|y_1\|^{-2}\|y_1\| + \|x_1\|\,\|y_1\|^{-2}\varepsilon + \|x_1\|\,\|x_1\|^{-4}(\|x_1\|\varepsilon + \|y_1\|\varepsilon)$$

$$\leqslant \varepsilon + 2\varepsilon + 3\varepsilon = 6\varepsilon. \qquad \blacksquare$$

We are now ready to prove the main theorem.

THEOREM 6. *Let*

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Then $\mathrm{dist}(\mathrm{Inv}\,A, \mathrm{Inv}\,B) < 1650\|A - B\|$ *for any* 3×3 *matrix* B *with* $\Omega_B = \{2;2,1;1,1;1\}$.

Proof. We first note that $P(\Omega_A)$ contains the following elements:

$$\Omega_A = \{1;2;2,1\},\ \{3;1,1,1;1;1;1\},\ \{1;1;3\},\ \{2;1,1;2,1\},$$

and $\{2;2,1;1,1;1\} = \Omega_B$.

We may assume with no loss of generality that $\|A - B\| = \varepsilon < \frac{1}{25}$. Since B has an eigenvalue λ with geometric multiplicity 2, we will have $\|(A - \lambda I) - (B - \lambda I)\| = \varepsilon$, and

$$B - \lambda I = \begin{bmatrix} \alpha' & c' & \beta' \\ \alpha'a' & c'a' & \beta'a' \\ \alpha'b' & c'b' & \beta'b' \end{bmatrix},$$

since $\dim \mathrm{Ker}(B - \lambda I) = 2$ implies that $B - \lambda I$ is a rank-one matrix. We observe that

$$A - \lambda I = \begin{bmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{bmatrix},$$

and since $\varepsilon < \frac{1}{25}$, we have that $\frac{25}{26} < |\delta| < \frac{25}{24}$, where $\delta = 1/c'$. We now have

$$C \equiv \delta(B - \lambda I) = \begin{bmatrix} \alpha & 1 & \beta \\ \alpha a & a & \beta a \\ \alpha b & b & \beta b \end{bmatrix}$$

and get the following inequalities:

$$|\beta| = |\delta \beta'| \leqslant \tfrac{25}{24}\varepsilon$$

as $|\beta'| \leqslant \|(A - \lambda I) - (B - \lambda I)\| = \varepsilon$,

$$|b| \leqslant \tfrac{25}{24}\varepsilon$$

similarly, and

$$|\lambda| = |\lambda + \beta' b' - \beta' b'| \leqslant |\lambda + \beta' b'| + |\beta' b'| \leqslant \varepsilon + |\beta b/\delta| \leqslant \varepsilon + \varepsilon = 2\varepsilon.$$

From these we now get

$$|a| = |\delta a' + \delta\lambda - \delta\lambda| \leqslant |\delta|\,|a' + \lambda| + |\delta\lambda| \leqslant \tfrac{25}{24}\varepsilon + \tfrac{50}{24}\varepsilon \leqslant 4\varepsilon,$$

$$|\alpha| = |\delta \alpha' + \delta\lambda - \delta\lambda| \leqslant |\delta|\,|\alpha' + \lambda| + |\delta\lambda| \leqslant \tfrac{25}{24}\varepsilon + \tfrac{50}{24}\varepsilon \leqslant 4\varepsilon.$$

We note that $\operatorname{Inv} C = \operatorname{Inv}(B - \lambda I) = \operatorname{Inv} B$ and $\operatorname{Inv}(A - \lambda I) = \operatorname{Inv} A$. The eigenspace of B corresponding to λ (or the $\operatorname{Ker} C$) is spanned by the two vectors

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -\alpha \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ -\beta \\ 1 \end{bmatrix}.$$

The eigenspace corresponding to the other eigenvalue of B (or the nonzero eigenvalue of C) is spanned by

$$\mathbf{x}_3 = \begin{bmatrix} 1 \\ a \\ b \end{bmatrix},$$

as the range of C is one-dimensional and spanned by this vector.

Now we would like to show there exists a constant k such that $\operatorname{dist}(\operatorname{Inv} A, \operatorname{Inv} B) \leqslant k\varepsilon$. To do this, we need to take any $M \in \operatorname{Inv} A$ and find $N \in \operatorname{Inv} B$ such that $\theta(M,N) \leqslant k\varepsilon$, and similarly we need to take any $N \in \operatorname{Inv} B$ and find $M \in \operatorname{Inv} A$ such that $\theta(M,N) \leqslant k\varepsilon$. To this end, we describe the invariant subspaces of A and B.

Let $M \in \operatorname{Inv} A$ be nontrivial (not $\{0\}$ or $\mathbb{C}^3$). Then $M = \operatorname{sp}\{t\mathbf{e}_1 + s\mathbf{e}_3\}$ or $M = \operatorname{sp}\{\mathbf{e}_1, t\mathbf{e}_2 + s\mathbf{e}_3\}$. If $M = \operatorname{sp}\{t\mathbf{e}_1 + s\mathbf{e}_3\}$, we may assume that $\|t\mathbf{e}_1 + s\mathbf{e}_3\| = 1$, so that $|t|, |s| \leqslant 1$. Let $N = \operatorname{sp}\{t\mathbf{x}_1 + s\mathbf{x}_2\}$. We note that $N \in \operatorname{Inv} B$ and we have

$$1 \leqslant \|t\mathbf{x}_1 + s\mathbf{x}_2\| = \sqrt{|t|^2 + |s|^2 + |t\alpha + s\beta|^2} \leqslant \sqrt{1 + (6\varepsilon)^2} \leqslant \sqrt{2}\,.$$

Also, $\|(t\mathbf{x}_1 + s\mathbf{x}_2) - (t\mathbf{e}_1 + s\mathbf{e}_3)\| \leqslant |t\alpha + t\beta| \leqslant 6\varepsilon$. So by Lemma 5, $\theta(M,N) \leqslant 6(6\varepsilon)$. If $M = \operatorname{sp}\{\mathbf{e}_1, t\mathbf{e}_2 + s\mathbf{e}_3\}$, where $\|t\mathbf{e}_2 + s\mathbf{e}_3\| = 1$, we know there are complex numbers u, v, and w such that $t\mathbf{e}_2 + s\mathbf{e}_3 = u\mathbf{x}_1 + v\mathbf{x}_2 + w\mathbf{x}_3$. Since $\langle \mathbf{e}_1, \mathbf{x}_3 \rangle = 1$, we must have that $w = 0$ if $u = 0$. Hence, if $u = 0$ we let $N = \operatorname{sp}\{\mathbf{x}_1, t\mathbf{e}_2 + s\mathbf{e}_3\} = \operatorname{sp}\{\mathbf{x}_1, \mathbf{x}_2\} \in \operatorname{Inv} B$. In this case $\|\mathbf{x}_1 - \mathbf{e}_1\| \leqslant |\alpha| \leqslant 4\varepsilon$ and $|\langle \mathbf{x}_1, t\mathbf{e}_2 + s\mathbf{e}_3 \rangle| = |t\alpha| \leqslant 4\varepsilon \leqslant \frac{1}{6}$. Applying Lemma 4, we get that

$$\theta(M,N) \leqslant \frac{272(4\varepsilon)}{\left(\frac{35}{36}\right)^2} \leqslant 340\varepsilon.$$

If $u \neq 0$, then $w \neq 0$ and we let $N = \operatorname{sp}\{\mathbf{x}_3, t\mathbf{e}_2 + s\mathbf{e}_3\} = \operatorname{sp}\{\mathbf{x}_3, u\mathbf{x}_1 + w\mathbf{x}_2\} \in \operatorname{Inv} B$. In this case $\|\mathbf{x}_3 - \mathbf{e}_1\| \leqslant \sqrt{|a|^2 + |b|^2} \leqslant 5\varepsilon$ and $|\langle \mathbf{x}_3, t\mathbf{e}_2 + s\mathbf{e}_3 \rangle| = |t^*a + s^*b| \leqslant 6\varepsilon < \frac{1}{4}$. Again, applying Lemma 4 we get that

$$\theta(M,N) \leqslant \frac{272(5\varepsilon)}{\left(\frac{15}{16}\right)^2} \leqslant 1650\varepsilon.$$

Now let $N \in \operatorname{Inv} B$ be nontrivial. Then $N = \operatorname{sp}\{t\mathbf{x}_1 + s\mathbf{x}_2\}$, $N = \operatorname{sp}\{\mathbf{x}_3\}$, $N = \operatorname{sp}\{\mathbf{x}_1, \mathbf{x}_2\}$ or $N = \operatorname{sp}\{t\mathbf{x}_1 + s\mathbf{x}_2, \mathbf{x}_3\}$. For $N = \operatorname{sp}\{t\mathbf{x}_1 + s\mathbf{x}_2\}$ we may assume that $|t|^2 + |s|^2 = 1$. Then $1 \leqslant \|t\mathbf{x}_1 + s\mathbf{x}_2\| \leqslant \sqrt{2}$. Let $M = \operatorname{sp}\{t\mathbf{e}_1 + s\mathbf{e}_3\}$

$\in \operatorname{Inv} A$. As in the previous paragraph, we get $\theta(M,N) \leqslant 6(6\varepsilon) = 36\varepsilon$. If $N = \operatorname{sp}\{\mathbf{x}_3\}$, then $M = \operatorname{sp}\{\mathbf{e}_1\} \in \operatorname{Inv} A$. We have

$$1 \leqslant \|\mathbf{x}_3\| \leqslant \sqrt{1 + |a|^2 + |b|^2} \leqslant \sqrt{1 + (4\varepsilon)^2 + (2\varepsilon)^2} \leqslant \sqrt{2}$$

and

$$\|\mathbf{x}_3 - \mathbf{e}_1\| = \sqrt{|a|^2 + |b|^2} \leqslant \sqrt{20\varepsilon^2} \leqslant 5\varepsilon.$$

Using Lemma 4, we get $\theta(M,N) \leqslant 6(5\varepsilon) = 30\varepsilon$. If $N = \operatorname{sp}\{\mathbf{x}_1, \mathbf{x}_2\}$, we let $M = \operatorname{sp}\{\mathbf{e}_1, \mathbf{e}_3\}$. Then $\|\mathbf{x}_1 - \mathbf{e}_1\| \leqslant |\alpha| \leqslant 4\varepsilon$, $\|\mathbf{x}_2 - \mathbf{e}_3\| \leqslant |\beta| \leqslant \frac{25}{24}\varepsilon$, $|\langle \mathbf{x}_1, \mathbf{x}_2 \rangle| \leqslant |\alpha\beta| \leqslant \frac{1}{4}\varepsilon$, and $|\langle \mathbf{e}_1, \mathbf{e}_3 \rangle| = 0$. So by results above on the norms of vectors and Lemma 4 we get that

$$\theta(N,M) \leqslant \frac{272(4\varepsilon)}{\left(\frac{35}{36}\right)^2} \leqslant 340\varepsilon.$$

Finally, suppose $N = \operatorname{sp}\{t\mathbf{x}_1 + s\mathbf{x}_2, \mathbf{x}_3\} = \operatorname{sp}\{t(\mathbf{x}_1 - \mathbf{x}_3) + s\mathbf{x}_2, \mathbf{x}_3\} = \operatorname{sp}\{\mathbf{y}, \mathbf{x}_3\}$, where $\mathbf{y} = w[t(\mathbf{x}_1 - \mathbf{x}_3) + s\mathbf{x}_2]$ and w is chosen so that $\|\mathbf{y}\| = 1$. We note that $\mathbf{y} = u\mathbf{e}_2 + v\mathbf{e}_3$ and $N = \operatorname{sp}\{\mathbf{y}, \mathbf{e}_1\} \in \operatorname{Inv} A$. Now $\|\mathbf{x}_3 - \mathbf{e}_1\| \leqslant 5\varepsilon$ from before, $\langle \mathbf{y}, \mathbf{e}_1 \rangle = 0$, and $|\langle \mathbf{y}, \mathbf{x}_3 \rangle| = |a^*u + b^*v| \leqslant |a| + |b| \leqslant \frac{9}{2}\varepsilon \leqslant \frac{1}{4}$. Another application of Lemma 4 yields $\theta(N,M) \leqslant 272(5\varepsilon)/(\frac{15}{16})^2 \leqslant 1650\varepsilon$. Hence, we have shown that

$$\begin{aligned} \operatorname{dist}(\operatorname{Inv} A, \operatorname{Inv} B) &\leqslant 1650\varepsilon \\ &\leqslant 1650\|A - B\| \qquad \text{when} \quad \|A - B| < \tfrac{1}{25}. \end{aligned}$$

However, this is also true when $\|A - B\| \geqslant \frac{1}{25}$, since $\operatorname{dist}(\operatorname{Inv} A, \operatorname{Inv} B) \leqslant 1$. ∎

REFERENCES

1 I. Gohberg and L. Rodman, On the distance between lattices of invariant subspaces, *Linear Algebra Appl.* 76:85–120 (1986).
2. I. Gohberg, P. Lancaster, and L. Rodman, *Invariant Subspaces of Matrices with Applications*, Wiley, New York, 1986.
3. R. Horn and C. Johnson, *Matrix Analysis*, Cambridge U.P., Cambridge, 1985.

Received 10 May 1989; final manuscript accepted 19 August 1989